Efficient Location Choice and the Returns to Agglomeration*

Sonia Jaffe† Stephen Morris‡

October 30, 2017

1 Introduction

Agents must make a binary choice about where to locate. They have heterogeneous preferences over locations. In addition, there are positive externalities from choosing the same location as others. Whichever location an agent chooses, his utility also includes an agglomeration effect that is increasing in the proportion of the population that chooses that location. Each individual is uncertain about the distribution of preferences in the population, but knows his own preferences, which serve as a noisy signal of population preferences. Under reasonable conditions,[1] there will be a unique equilibrium where the majority of agents go to the more popular location, but agents with sufficiently extreme preferences go to the other location. However, agglomeration externalities imply that there may be over- or under-agglomeration in equilibrium relative to the efficient outcome.

We show that if the marginal returns to agglomeration are sufficiently decreasing in the proportion of the population co-locating (returns are more concave than $\log(\cdot)$) then there will be over-agglomeration. The marginal (indifferent) person generates a much larger marginal benefit at the smaller location, so even though fewer people benefit, it would be social efficient for them to move there, thereby decreasing agglomeration. If the agglomeration function is less concave than $\log(\cdot)$ (for example, linear) there will be under-agglomeration in equilibrium. The marginal (indifferent) person benefits more people if they go to the location with more people. If the marginal social benefit at that location is not too much smaller

*We thank Rosa Argenziano and Scott Kominers for helpful comments. This work grew out of related joint work with Scott. We also thank Evgeniya Lambrinaki for valuable research assistance.

†University of Chicago spj@uchicago.edu. Jaffe gratefully acknowledges the support of a NSF Graduate Research Fellowship and a Terence M. Considine Fellowship in Law and Economics.

‡Princeton University smorris@princeton.edu

[1] It is sufficient that either there is sufficient homogeneity in preferences or there is sufficient heterogeneity; see discussion in Section 4.1.

than at the location with fewer people, it is socially efficient to move the marginal person to the location with more people, thereby increasing agglomeration. Analogous results hold then there are multiple equilibria, idiosyncratic signals of common preferences over locations and negative externalities of co-locating.

The model has natural interpretations in a variety of contexts and the critical condition on the concavity of agglomeration thus has corresponding economic interpretations. If firms are making decisions about which city to locate in, agglomeration may correspond to positive spillovers via the labor market, product market, or infrastructure. If high quality students are making decisions about which university to attend, agglomeration might reflect the benefit to high quality students locating together. If individuals are deciding between alternative technologies or standards, agglomeration corresponds to benefit to standardization. In each case, there is a natural interpretation to the concavity of the agglomeration, i.e., the rate at which the marginal benefit of agglomeration falls as the level of agglomeration increases.

The conflict between private and social benefits of agglomeration are studied in Mitchell and Skrzypacz (2006) and Argenziano (2008).[2] In a different, dynamic, model, Mitchell and Skrzypacz (2006) use the concavity of $\ln(\cdot)$ condition.[3] Our model is the static one of Argenziano (2008); we thus extend her result showing that there is under-agglomeration with linear agglomeration and normally distributed benefits. As we discussion in Section 4.5, our paper also connects to the empirical literature estimating agglomeration effects of firms co-locating, especially the more recent efforts to capture non-linearities of these spillovers.

2 Model

There are two sectors and a unit mass of agents. If proportion l of the population is in sector 1, then agent i gets utility $\frac{1}{2}x_i + h(l)$ from being in sector 1 and $-\frac{1}{2}x_i + h(1-l)$ from being in sector 2, where h is an increasing function, bounded on the interval $(0,1)$ by $\bar{h}$. The difference in agglomeration benefits between the two sectors is

$$\tilde{h}(l) = h(l) - h(1-l).$$

Observe that $\tilde{h}$ is increasing, $\tilde{h}\left(\frac{1}{2}\right) = 0$, and $\tilde{h}$ has rotational symmetry around $\frac{1}{2}$, i.e., $\tilde{h}(l) = -\tilde{h}(1-l)$ for all l. An individual's decision depends on the difference in utility between sector 1 and sector 2, $x_i + \tilde{h}(l)$.

[2] Both papers are about the impact of endogenizing prices in such settings, but address the problem with exogenous prices first.

[3] They assume (page 326) that (the level of) agglomeration times the marginal benefit of agglomeration is increasing in agglomeration, a condition equivalent to the $\ln(\cdot)$ condition.

A state θ is distributed according to a symmetric, unimodal, continuous prior distribution g with mean μ, which we assume w.l.o.g is greater than zero. Each agent's benefit is $x_i = \theta + \varepsilon_i$, where ε_i is distributed according to a symmetric density f, with mean zero and full support. After observing x_i, agent i believes θ is distributed according to the posterior $g(\theta|x_i)$. We assume benefits satisfy *first order stochastic dominance* (FOSD): observing a higher x_i leads to a FOSD increase in beliefs about θ, that is

$$\frac{\partial}{\partial x}\int_{-\infty}^{k} g(\theta|x_i)d\theta < 0 \qquad\qquad \forall k, x_i.$$

The FOSD property guarantees that any equilibrium is in monotone strategies, where an agent goes to sector 1 if and only if his benefit x_i is greater than a threshold, $\hat{x}$. Moreover, since there is an increased incentive to locate in sector 1 when others locate in sector 1 and when the benefit x_i is higher, the game is "monotone supermodular" (in the language of Van Zandt and Vives (2007)). Monotone supermodularity implies the existence of largest and smallest monotone pure strategy equilibria. Because the game is symmetric across agents, these equilibria will be symmetric.

A player who has benefit x_i and thinks others follow a strategy with threshold $\hat{x}$ gets payoff

$$U(x_i, \hat{x}) = x_i + \int_{-\infty}^{\infty} \tilde{h}\left(1 - F(\hat{x} - \theta)\right) g(\theta|x_i)\, d\theta$$

for sector 1 relative to sector 2. A threshold $\hat{x}$ is an equilibrium if $U(\hat{x}, \hat{x}) = 0$. To guarantee that the equilibrium is unique, we maintain the assumption that payoffs are *single-crossing:* that is, $U(x, x) = 0$ has a unique solution. A sufficient condition for this maintained assumption is that heterogeneity is sufficiently small, so that f is sufficiently concentrated. We discuss this and other related sufficient conditions on primitives for single-crossing, in Section 4.1.

Because $U(x, x)$ is continuous, $U(\mu, \mu) = \mu > 0$ and $U(-\bar{h} - \varepsilon, -\bar{h} - \varepsilon) < 0$, there will always be an equilibrium at some threshold $\hat{x} < \mu$.

3 Agglomeration and Welfare

It is immediate that social welfare is maximized by allocating agents according to a threshold strategy. If an agent allocated to sector 2 has a lower benefit than an agent allocated to sector 1, welfare is increased by swapping them. We will say there is over-agglomeration, when the equilibrium threshold implies a less even distribution of people between the two sectors than the socially optimal threshold. Since we have assumed $\mu > 0$, this corresponds to *there is over (under) agglomeration if the equilibrium threshold is below (above) the socially optimal threshold.*

Our main result is:

Proposition 1. *When h is more (less) concave than $\log(\cdot)$, there is over (under) agglomeration.*

Suppose everyone uses the threshold strategy of going to sector 1 if and only if $x_i \geq \hat{x}$. Ex-ante welfare (for an arbitrary cutoff) is

$$W(\hat{x}) = \int_{-\infty}^{\infty} \left[\begin{array}{l} \frac{1}{2} \int_{\varepsilon=\hat{x}-\theta}^{\infty} (\theta+\varepsilon) f(\varepsilon) \, d\varepsilon - \frac{1}{2} \int_{\varepsilon=-\infty}^{\hat{x}-\theta} (\theta+\varepsilon) f(\varepsilon) \, d\varepsilon \\ +(1-F(\hat{x}-\theta)) h(1-F(\hat{x}-\theta)) + F(\hat{x}-\theta) h(F(\hat{x}-\theta)) \end{array} \right] g(\theta) \, d\theta$$

Differentiating with respect to $\hat{x}$ gives

$$\begin{aligned} W'(\hat{x}) &= \int_{-\infty}^{\infty} \left[\begin{array}{l} -\hat{x} f(\hat{x}-\theta) - f(\hat{x}-\theta)(1-F(\hat{x}-\theta)) h'(1-F(\hat{x}-\theta)) \\ -f(\hat{x}-\theta) h(1-F(\hat{x}-\theta)) + f(\hat{x}-\theta) h(F(\hat{x}-\theta)) \\ +f(\hat{x}-\theta) F(\hat{x}-\theta) h'(F(\hat{x}-\theta)) \end{array} \right] g(\theta) \, d\theta \\ &= W_1' + W_2', \end{aligned}$$

where

$$W_1' = \int_{-\infty}^{\infty} \left(-\hat{x} + h(F(\hat{x}-\theta)) - h(1-F(\hat{x}-\theta))\right) f(\hat{x}-\theta) g(\theta) \, d\theta$$

$$W_2' = \int_{-\infty}^{\infty} \left(F(\hat{x}-\theta) h'(F(\hat{x}-\theta)) - (1-F(\hat{x}-\theta)) h'(1-F(\hat{x}-\theta))\right) f(\hat{x}-\theta) g(\theta) \, d\theta$$

The first part of the welfare derivative (W_1') is the value to the marginal person of switching sectors when the threshold moves. If the threshold increases they lose $\hat{x}$ in private value from leaving sector 1 and gain or lose $h(F(\hat{x}-\theta)) - h(1-F(\hat{x}-\theta))$ from the agglomeration difference. This is proportional to the difference in utility between the same sectors for the marginal person, that is $W_1' \propto -U(x,x)$.[4] We therefore know that W_1' is zero at the equilibrium $\tilde{x}$ and $W_1'(x) > 0$ if and only if $x < \tilde{x}$.

[4]This uses $g(\theta|x_i) \propto f(\hat{x}-\theta) g(\theta)$, they are not equal because one is conditional on being the marginal person and the other includes the probability of being the marginal person.

The second part of the welfare derivative (W_2') is the effect of changing the threshold on everyone's value from agglomeration. It has rotational symmetry around μ, that is $W_2'(\mu + \Delta) = -W_2'(\mu - \Delta)$. For any threshold $\hat{x} < \mu$, more weight will be where $\theta > \hat{x}$ and $F(\hat{x} - \theta) < 1 - F(\hat{x} - \theta)$. This means that $W_2'(\hat{x}) > 0$ for $\hat{x} < \mu$ if and only if $F(a)h'(F(a)))$ is decreasing, which corresponds to $h(\cdot)$ being more concave than $\log(\cdot)$. If h is more concave than $\log(\cdot)$ then for $\hat{x} < \tilde{x}$ both W_1' and W_2' are positive and if $\hat{x} > \mu$ both W_1' and W_2' are negative; in either case their sum cannot be zero. So the social optimum must lie in $(\tilde{x}, \mu)$. Moving the threshold towards the mean decreases agglomeration, so we see that when h' is more concave than $\log(\cdot)$, the social optimum has less agglomeration than the global games equilibrium.

When $h(\cdot)$ is less concave than $\log(\cdot)$, then for $\hat{x} \in (\tilde{x}, \mu)$ both W_1' and W_2' are negative, so the optimum cannot be in this range. It is sufficient to show that in this case the optimum cannot be at a threshold above the mean. To see this, we return to the welfare function and, for any $x > \mu$, we look at the symmetric point x' such that $x - \mu = \mu - x' = \Delta$ and show that welfare is higher at x' that at x. Because they result in the same distribution of agglomeration,[5] the difference in welfare between the two thresholds is

$$W(\mu - \Delta) - W(\mu + \Delta) = \int_{-\infty}^{\infty} \left(\int_{\mu-\Delta-\theta}^{\mu+\Delta-\theta} (\theta + \varepsilon) f(\varepsilon) \, d\varepsilon \right) g(\theta) \, d\theta,$$

which is zero at $\Delta = 0$ and

$$\frac{\partial}{\partial \Delta} \left(W(\mu - \Delta) - W(\mu + \Delta) \right) = \int_{-\infty}^{\infty} \left((\mu + \Delta) f(\mu + \Delta - \theta) + (\mu - \Delta) f(\mu - \Delta - \theta) \right) g(\theta) \, d\theta$$

$$= \int_{-\infty}^{\infty}$$

The first term is positive and the second term is zero because $g(\theta)$ is symmetric around μ.[6] So for $\Delta > 0$, we have $W(\mu - \Delta) - W(\mu + \Delta) > 0$, implying that the globally optimal threshold cannot be greater than μ. Therefore the optimal threshold is less than the equilibrium

[5] Argenziano (2008) shows this for normal distributions. By taking the derivative with respect to Δ we do not need to rely on properties of the normal distribution, just the symmetry of the distributions.

[6] Symmetry around μ implies that $g(\theta) = g(2\mu - \theta)$ and

$$\begin{aligned} f(\mu + \Delta - \theta) - f(\mu - \Delta - \theta) &= (f(-\mu - \Delta + \theta) - f(-\mu + \Delta + \theta) \\ &= -\left(f(\mu + \Delta - (2\mu - \theta)) - f(\mu - \Delta - (2\mu - \theta)) \right) \end{aligned}$$

so the overall integral is zero.

threshold $\tilde{x}$; since the equilibrium threshold is below the mean ($x^* < \tilde{x} < \mu$), this means the optimal threshold results in more agglomeration than the global games equilibrium.

4 Discussion

4.1 Uniqueness

We now discuss primitive assumptions under which the maintained single-crossing assumption holds.

First, consider the case where the agglomeration function is linear, with $h(l) = \frac{1}{2}l$ and thus $\tilde{h}(l) = l - \frac{1}{2}$; and the benefits are normally distributed, with $\theta \sim N(\mu, \frac{1}{\alpha})$ and $\varepsilon_i \sim N(0, \frac{1}{\beta})$. In this case, one can confirm by simple calculation that single-crossing is satisfied if

$$\frac{\alpha^2 \beta}{(\alpha + \beta)(\alpha + 2\beta)} \leq 2\pi.$$

The calculations appear in Morris and Shin (2005) and this corresponds to the case analyzed in Argenziano (2008).

Notice that this condition is automatically satisfied if we fix α (the precision of θ) and let β (the precision of the idiosyncratic component) tend to infinite. This corresponds to looking at what happens in the limit as heterogeneity disappears. In fact, there is a unique equilibrium in general as heterogeneity disappears. This is a key finding of the global games literature (Carlsson and Van Damme, 1993; Morris and Shin, 2003).[7]

The uniqueness condition for the linear normal case is also automatically satisfied if we fix β and let α tend to infinite. This corresponds to looking at what happens in the limit as heterogeneity blows up. In this case, there is a unique equilibrium because the benefits terms swamps the strategic effect. Morris and Shin (2005) show that there will be uniqueness in general as heterogeneity blows up as long as the derivative of the agglomeration function h is bounded above and below. They also discuss the relation to uniqueness sufficient conditions based on large heterogeneity in the literature.

4.2 With multiple equilibria

When there are multiple equilibria, we do not know the shape of the function $U(x,x)$ (the utility of threshold-person); therefore we cannot globally sign the first part of the welfare

[7] This literature has focused on the "common value" case where the idiosyncratic component of agents' types are payoff-relevant, but the results also apply when in the "private value" case analyzed here where the idiosyncratic component is payoff-relevant. This common value / private value comparison is discussed in Morris and Shin (2005). Note that first order stochastic dominance is not needed as an assumption in this literature because it is automatically satisfied in the limit as heterogeneity disappears.

derivative, W_1'. However, we still know that at any threshold equilibrium $W_1'(\tilde{x}) = 0$, so we can sign the local derivative of welfare with respect to the threshold and therefore with respect to agglomeration. At a threshold equilibrium, welfare will be *locally* increasing in agglomeration if and only if h' is less concave than $\log(\cdot)$.

4.3 Common Value

In an alternative common value model, $x_i = \theta + \epsilon_i$, is not a private value, but a private signal about the common true value, θ. The expected value of agglomeration is unchanged, but the individual cares about $E[\theta|x]$ for the direct part of their payoff. As long as the social planner and individual agree that θ is the true value, the conditions for over and under-agglomeration are unaffected. The derivative of welfare still breaks down into the the marginal person's utility and the effect via the agglomeration benefits, which depends on $\frac{\partial}{\partial a}(ah'(a)) > 0$.

If the planner cares about θ and individuals care about x, there is an additional term in the derivative of the planner's welfare function, $(\hat{x} - E[\theta|\hat{x}])Pr[\hat{x}]$, which always pushes for more agglomeration. The planner thinks the marginal person is less different from the average than the person themself does; so what is right for the majority of people is more likely to be right for the marginal person. The previous condition $\frac{\partial}{\partial a}(ah'(a)) > 0$ is now sufficient, but not necessary for under agglomeration.

4.4 Negative externality.

There may be reasons – traffic, social signaling, and limited resources – for people to want to go to the location with fewer people. In our model this corresponds to $h' < 0$. This makes uniqueness less likely,[8] but conditional on a unique threshold equilibrium existing, the same conditions for over- and under-agglomeration hold. If $h'' > 0$ then the marginal person at the larger location does less harm than at the smaller location; if the effect is enough smaller ($h'' > -ah(a) \Rightarrow \frac{\partial}{\partial a}(ah'(a)) > 0$) than the total damage is less even though the harm is inflicted on more people. Conversely, if $h'' < -ah(a)$ then the fact that the harm is done to more people outweighs the lower level of harm and the social planner would like to decrease agglomeration.

4.5 Empirical Concavity

While there is a substantial literature that tries to estimate agglomeration effects and spillovers in location choice, those with sufficient data to look at the curvature (as op-

[8]With $h' < 0$, first-order stochastic dominance no longer ensures that every equilibrium is a threshold equilibrium. There may be an equilibrium where person A with x_A does not want to go to sector 1 because she thinks there will be a lot of people there, but person B with signal $x_B < x_A$, will go to sector one (even though she likes it less than person A) because she does not think there will be many people in sector 1.

posed to the slope) are more limited.[9] Martin et al. (2011) analyze plant level data and find agglomeration spillovers on productivity that are bell-shaped, suggesting some areas are overly concentrated and others are under-concentrated. Maré and Graham (2013) find decreasing returns to agglomeration for most industries in New Zealand, but do not report the parameters to check whether the effects are more concave than $\log(\cdot)$. Davis and Henderson (2008) look at the location of firms' headquarters in the United States. The estimates from their quadratic specification suggest that the spillovers are more concave than $\log(\cdot)$ once there are more than 7 headquarters in an area, meaning they are generally overly concentrated. Cainelli et al. (2015) also find spillovers are more concave than $\log(\cdot)$ using a cubic specification.

References

Argenziano, R. (2008). Differentiated networks: Equilibrium and efficiency. *The RAND Journal of Economics*, 39(3):747–769.

Cainelli, G., Fracasso, A., and Vittucci Marzetti, G. (2015). Spatial agglomeration and productivity in italy: a panel smooth transition regression approach. *Papers in Regional Science*, 94(S1):S39–S67.

Carlsson, H. and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, pages 989–1018.

Davis, J. C. and Henderson, J. V. (2008). The agglomeration of headquarters. *Regional Science and Urban Economics*, 38(5):445–460.

Maré, D. C. and Graham, D. J. (2013). Agglomeration elasticities and firm heterogeneity. *Journal of Urban Economics*, 75:44–56.

Martin, P., Mayer, T., and Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. *Journal of Urban Economics*, 69(2):182–195.

Mitchell, M. F. and Skrzypacz, A. (2006). Network externalities and long-run market shares. *Economic Theory*, 29(3):621–648.

Morris, S. and Shin, H. S. (2003). Global games: Theory and applications. In Hansen;, M. D. L. P. and Turnovsky, S. J., editors, *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge University Press, Cambridge.

Morris, S. and Shin, H. S. (2005). Heterogeneity and uniqueness in interaction games. In Blume, L. and Durlauf, S., editors, *The Economy as an Evolving Complex System III*. Oxford University Press.

Sokullu, S. (2016). Network effects in the german magazine industry. *Economics Letters*, 143:77–79.

Van Zandt, T. and Vives, X. (2007). Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory*, 134(1):339–360.

[9] Sokullu (2016) estimate nonlinear network externalities in the German magazine industry.